AD-A045 842 MASSACHUSETTS INST OF TECH CAMBRIDGE ELECTRONIC SYST--ETC F/G 12/1

THE USE OF HARMONIC ANALYSIS IN SUBOPTIMAL ESTIMATOR DESIGN.(U)

SEP 77 S I MARCUS, A S WILLSKY, K HSU AFOSR-72-2273

UNCLASSIFIED AFOSR-TR-77-1250 NL

| OF |

AD A045842

END

DATE FILMED

11 - 77

DDC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR-77-1250	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE USE OF HARMONIC ANALYSIS IN SUBOPTIMAL ESTIMATOR DESIGN		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Steven I. Marcus Alan S. Willsky Kai Hsu		8. CONTRACT OR GRANT NUMBER(s) AFOSR-72-2273 NSF-GK-42090
9. PERFORMING ORGANIZATION NAME AND ADDRESS Massachusetts Institute of Technology Electronic Systems Laboratory Cambridge, Massachusetts 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB DC 20332		12. REPORT DATE Sep 77
		13. NUMBER OF PAGES 28
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 29 p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DDC
RECEIVED
NOV 2 1977
F

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The state estimation problem for bilinear stochastic systems evolving on compact Lie groups and homogeneous spaces is considered. The problem is motivated by some applications involving rotational processes in three dimensions. The theory of harmonic analysis on compact Lie groups is used to define assumed density approximations which result in implementable suboptimal estimators for the state of → next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 127200 mt

AD A045842

DDC FILE COPY

20. Abstract (continued)

the bilinear system. The results of Monte Carlo simulations are reported; these indicate that simple filters designed by these techniques perform well as compared to other filters.

THE USE OF HARMONIC ANALYSIS IN SUBOPTIMAL ESTIMATOR DESIGN*

Steven I. Marcus
Department of Electrical Engineering
The University of Texas at Austin
Austin, Texas 78712

Alan S. Willsky
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, Mass. 02139

Kai Hsu
Department of Electrical Engineering
The University of Texas at Austin
Austin, Texas 78712

Abstract

The state estimation problem for bilinear stochastic systems evolving on compact Lie groups and homogeneous spaces is considered. The problem is motivated by some applications involving rotational processes in three dimensions. The theory of harmonic analysis on compact Lie groups is used to define assumed density approximations which result in implementable suboptimal estimators for the state of the bilinear system. The results of Monte Carlo simulations are reported; these indicate that simple filters designed by these techniques perform well as compared to other filters.

*This work was supported by NSF under Grants GK-42090 and ENG76-11106, by AFOSR under Grant 72-2273, by an NSF Fellowship (S.I. Marcus) and by the Joint Services Electronics Program under Contract F44620-71-C-0091. It is based in part on the Ph.D. dissertation by S. Marcus submitted to the Massachusetts Institute of Technology, Cambridge, Mass.

I. INTRODUCTION

Fourier series analysis has been applied in several recent studies [1] - [4] to estimation problems for stochastic processes evolving on the circle S^1. Willsky [4] used Fourier series methods to define "assumed density" approximations for certain phase tracking and demodulation problems. In fact, a system designed using these techniques performed better than other estimators, including an optimal phase-lock loop.

In this paper we study bilinear systems evolving on compact Lie groups or homogeneous spaces [29]. The optimal estimator is in general infinite dimensional [7], and our approach to the design of suboptimal estimators is a generalization of that of Willsky [4], whose work is reviewed briefly in Section III. The basic approach involves the definition of an "assumed density" form for the conditional density of the system state at time t given observations up to time t. These densities are defined via the techniques of harmonic analysis on compact Lie groups [5],[6] (which generalize the Fourier series on the Lie group S^1). Our method differs from most previous assumed density approximations in that our approximation is defined on the appropriate compact manifold (as opposed to the usual Gaussian approximations, for example, which are defined on $\mathbb{R}^n$ [7]). This method also avoids the problem of merely truncating higher order terms in a harmonic expansion; as pointed out by Lo [18] and Willsky [4], such higher order terms will not be negligible, especially if the filter is performing well. For an alternative approach to discrete-time estimation problems on Lie groups and homogeneous spaces, see the work of Lo and Eshleman [18-20], who use exponential Fourier densities to avoid the truncation problem.

In Section II we review some general properties of stochastic bilinear systems and discuss the estimation problem for systems evolving on compact Lie groups and homogeneous spaces. Section III contains the application of the technique to systems evolving on S^n, while Section IV contains the application to systems on SO(n). Results of Monte Carlo simulations of the S^2 estimator are presented in Section V. Finally, the Appendix reviews some necessary concepts from the theory of harmonic analysis on compact Lie groups. Some of the concepts of this paper were introduced in [21] and [22], but no simulation results were presented.

II. ESTIMATION FOR STOCHASTIC BILINEAR SYSTEMS

The basic stochastic bilinear system (or linear system with state-dependent noise) considered here is described by the Ito stochastic differential equation [4], [9-17], [21], [25]

$$dx(t) = \{[A_0 + \tfrac{1}{2} \sum_{i,j=1}^{N} Q_{ij}(t) A_i A_j]dt + \sum_{i=1}^{N} A_i dw_i(t)\} x(t) \quad (1)$$

where x is an n-vector or an n x n matrix, the A_i are n x n matrices, Q_{ij} is the (i,j)th element of Q, and w is a Brownian motion (Wiener) process with strength Q(t) such that $E[w(t)w'(s)] = \int_0^{\min(t,s)} Q(\tau)d\tau$. Following the notation of [8-11], we define $\mathscr{L} = \{A_0, A_1, \ldots, A_N\}_{LA}$ to be the smallest Lie algebra containing these matrices. The corresponding connected matrix Lie group is denoted by $G = \{\exp\mathscr{L}\}_G$. Then, if x is an n x n matrix and $x(t_0) \epsilon G$, the solution x(t) of (1) evolves on G (i.e., $x(t) \epsilon G$ for all $t \geq 0$) in the mean-square sense and almost surely [15-17]. If x is an n-vector, then the solution of (1) evolves on the homogeneous space $G \cdot x(t_0)$.

Associated with the Ito equation (1) is a sequence of equations for the powers of the state x(t) (see Brockett [9], [10]). If N(n,p) denotes the binomial coefficient $\binom{n+p-1}{p}$, then given an n-vector x, we define $x^{[p]}$ to be the N(n,p)-vector with components equal to the monomials (homogeneous polynomials) of degree p in $x_1, \ldots, x_n$, the components of x, scaled so that $\|x\|^p = \|x^{[p]}\|$. Given an m x n matrix A, we denote by $A^{[p]}$ the unique matrix which verifies

$$y = Ax \Rightarrow y^{[p]} = A^{[p]} x^{[p]} . \quad (2)$$

$A^{[p]}$ can be interpreted as a linear operator on symmetric tensors of degree p [9], and is known as the symmetrized Kronecker p^{th} power of A [26]. It is clear that if x satisfies the linear differential equation

$$\dot{x}(t) = Ax(t) \tag{3}$$

then $x^{[p]}$ also satisfies a linear differential equation

$$\dot{x}^{[p]}(t) = A_{[p]}x^{[p]}(t) \quad . \tag{4}$$

We regard this as the definition of $A_{[p]}$, which is the infinitesimal version of $A^{[p]}$. In fact, $A_{[p]}$ can be easily computed from A [25].

It can easily be shown that if x satisfies (1), then $x^{[p]}$ satisfies the Ito equation

$$dx^{[p]}(t) = \{A_{0_{[p]}} + \tfrac{1}{2}\sum_{i,j=1}^{N} Q_{ij}(t)A_{i_{[p]}}A_{j_{[p]}}\}x^{[p]}(t)dt + \sum_{i=1}^{N} A_{i_{[p]}}x^{[p]}(t)dw_i(t) \tag{5}$$

In addition, if the nxn matrix X satisfies (1), it is easy to show that $X^{[p]}$ also satisfies (5). As we shall see later in the section, this sequence of equations is a valuable tool in the study of state estimation.

The observation model considered in this paper consists of linear observations of the state corrupted by additive white noise, or

$$dz(t) = L(x(t))\,dt + dv(t) \tag{6}$$

where L is a linear operator and v is a Wiener process. This bilinear system-linear observation model is useful in the study of certain practical problems, such as the S^2 satellite tracking and SO(3) rigid body orientation estimation problems discussed in [13, Ch. 4] and [21, Sec. IV].

The remainder of the paper is devoted to the study of the estimation problem for two classes of systems of the form (1), (6), which are motivated by the aforementioned examples. The first system consists of the bilinear state equation

$$dX(t) = [A_0 + \tfrac{1}{2} \sum_{i,j=1}^{N} Q_{ij}(t) A_i A_j] X(t) dt + \sum_{i=1}^{N} A_i X(t) dw_i(t) \tag{7}$$

with linear measurements

$$dz_1(t) = X(t) h(t) dt + R^{\frac{1}{2}}(t) dv(t) \tag{8}$$

where $X(t)$ and $\{A_i\}$ are nxn matrices, $z_1(t)$ is a p-vector, w is a Wiener process with strength $Q(t) \geq 0$, v is a standard Wiener process independent of w, and $R > 0$. More general linear measurements can obviously be considered, but for simplicity of notation we restrict our attention to (8), which arises in the star tracking example of [13, Ch. 4]. We will make the crucial assumption that the Lie group $G = \{\exp \mathscr{L}\}_G$ is compact.

The second system consists of the bilinear state equation (1) with linear measurements

$$dz_2(t) = H(t) x(t) dt + R^{\frac{1}{2}}(t) dv(t) \tag{9}$$

where $x(t)$ is an n-vector, A_i are n x n matrices, and z_2, v and w are as above. It will be assumed that x evolves on a compact homogeneous space [8], [13], [29]--i.e., the solution of (1) is

$$x(t) = X(t) x(0) \tag{10}$$

where X satisfies (7) with $X(0) = I$ and evolves on the compact Lie group $G = \{\exp \mathscr{L}\}_G$.

It is shown in [13] that, by a linear change of basis on the state space, (7) and (1) can be transformed into equations which evolve on the compact special orthogonal group $SO(n) \triangleq \{X \in \mathbb{R}^{n \times n} | X'X = I, \det X = +1\}$ and the compact homogeneous space $S^{n-1} = \{x \in \mathbb{R}^n | x'x = 1\}$ (the (n-1)-sphere), respectively. Hence, we need only consider systems evolving on SO(n) and S^{n-1} (this is equivalent to the assumption that $A_0, A_1, \ldots, A_N$ are skew-symmetric [13]).

The estimation criterion which will be used for these two problems is the constrained least-squares estimator, which is analogous to the criterion used in [1], [4], and [21] for the phase estimation problem. That is, for (7)-(8) we wish to find $\tilde{X}(t|t)$ which minimizes the conditional error covariance

$$J_1 = E[tr\{(X(t) - \tilde{X}(t|t))'(X(t) - \tilde{X}(t|t))\}|z_1^t] \tag{11}$$

subject to the SO(n) constraint $\tilde{X}(t|t)'\tilde{X}(t|t) = I$, where the notation (11) denotes the conditional expectation given the σ-field $\sigma\{z_1^t\}$ generated by the observed process $z_1^t \triangleq \{z_1(s), 0 \leq s \leq t\}$ up to time t. For (1), (9) we seek $\tilde{x}(t|t)$ which minimizes

$$J_2 = E[(x(t) - \tilde{x}(t|t))' (x(t) - \tilde{x}(t|t))|z_2^t] \tag{12}$$

subject to the S^{n-1} constraint $\|\tilde{x}(t|t)\|^2 = \tilde{x}(t|t)'\tilde{x}(t|t) = 1$.

It is easily shown [13] that the optimal estimates are, respectively

$$\tilde{X}(t|t) = \pm\hat{X}(t|t)[\hat{X}(t|t)'\hat{X}(t|t)]^{-\frac{1}{2}} \tag{13}$$

$$\tilde{x}(t|t) = \frac{\hat{x}(t|t)}{\|\hat{x}(t|t)\|} \tag{14}$$

where the conditional expectation is denoted by the equivalent notations

$$\hat{x}(t|t) \triangleq E[x(t)|z_2^t] \triangleq E^t[x(t)] \tag{15}$$

The sign in (13) is chosen to insure that det $\tilde{X}(t|t) = +1$ [23]. Thus in both cases we must compute the conditional expectation of the state (X(t) or x(t)) given the past observations ($z_1^t = \{z_1(s), 0 \leq s \leq t\}$ or $\{z_2^t(s), 0 \leq s \leq t\}$).

The equations for computing the conditional expectation can be derived from the general nonlinear filtering equation [7] and the moment equation (5). The resultant equations for the SO(n) system (7)-(8) are

$$dE^t[X_v^{[p]}(t)] = [(A_{0_{[p]}} + \tfrac{1}{2}\sum_{i,j=1}^{N} Q_{ij}(t)A_{i_{[p]}}A_{j_{[p]}}) \otimes I]\, E^t[X_v^{[p]}(t)]dt$$

$$+ \{E^t[X_v^{[p]}(t)h'(t)X(t)] - E^t[X_v^{[p]}(t)]h'(t)E^t[X(t)]\}R^{-1}(t)d\nu_1(t) \qquad (16)$$

$$d\nu_1(t) = dz_1(t) - \hat{X}(t|t)h(t)dt \qquad (17)$$

where $\otimes$ denotes Kronecker product and $X_v^{[p]}$ is the vector containing the elements of the matrix $X^{[p]}$ in lexicographic order [26], [32, p.64]. For the S^{n-1} system (1), (9), we have

$$dE^t[x^{[p]}(t)] = [A_{0_{[p]}} + \tfrac{1}{2}\sum_{i,j=1}^{N} Q_{ij}(t)A_{i_{[p]}}A_{j_{[p]}}]E^t[x^{[p]}(t)]dt$$

$$+ \{E^t[x^{[p]}(t)x'(t)] - E^t[x^{[p]}(t)]E^t[x'(t)]\}H'(t)R^{-1}(t)d\nu_2(t) \qquad (18)$$

$$d\nu_2(t) = dz_2(t) - H(t)\hat{x}(t|t)dt\ . \qquad (19)$$

The structure of these equations is quite similar to that of [4]--i.e., each estimator consists of an infinite band of filters, and the filter for the p^{th} moment is coupled only to those for the first and $(p+1)^{st}$ moments. Therefore, we are led to the design of suboptimal estimators. The technique proposed here is motivated by the highly successful use of folded normal assumed density approximations in the phase tracking problem [4]; filters designed using this technique performed very well as compared with other suboptimal estimators. We will describe similar techniques for the design of suboptimal estimators on S^n and SO(n).

We first review the notions of Brownian motion and Gaussian densities on Lie groups and homogeneous spaces. Yosida [28] proved that the fundamental solution of

$$\frac{\partial p(x,t)}{\partial t} - \gamma \Delta p(x,t) = 0 \tag{20}$$

where $\gamma > 0$ and Δ is the Laplace-Beltrami operator (Laplacian) on a Riemannian homogeneous space M [5,13,29], is the density (with respect to the Riemannian measure) of a Brownian motion on M [1].

According to [5], the fundamental solution of (20) is given by

$$p(x,t;x_0,t_0) = \sum_i \phi_i(x)\phi_i(x_0)e^{-\lambda_i(t-t_0)\gamma} \tag{21}$$

where λ_i and φ_i are the eigenvalues and the corresponding eigenfunctions of the Laplacian (see the Appendix). The function $p(x,t;\, x_0,t_0)$ is the solution to (20) with initial condition equal to the singular distribution concentrated at $x = x_0$. Also, Grenander [27] defines a Gaussian (normal) density to be the solution of (20) for some t.

The folded normal density $F(\theta;\eta,\gamma)$ used by Willsky as an assumed density approximation for the phase tracking problem is indeed a normal density on S^1 in the sense of Grenander [4]; in fact, the trigonometric polynomials $e^{-in\theta}$ are eigenfunctions of the Laplacian on S^1. Motivated by the success of Willsky's suboptimal filter, we will design suboptimal estimators for the SO(n) and S^n bilinear systems by employing normal assumed

---

[1] Yosida defines a Brownian motion process to be a temporally and spatially homogeneous Markov process on M which satisfies a continuity condition of Lindeberg's type.

conditional densities of the form

$$p(x,t) = \sum_i \phi_i(x)\phi_i(\eta(t))e^{-\lambda_i\gamma(t)} \tag{22}$$

where $\eta(t)$ and $\gamma(t)$ are parameters of the density which are to be estimated.[2]

[2] In order to assure the existence of a conditional density, it is sufficient to assume that the system is "controllable from the noise" [10,13,17].

III. ESTIMATION ON S^n

In this section the suboptimal estimation technique discussed in the previous section will be used in order to design filters for the S^n estimation problem (1),(9). The optimal constrained least-squares estimator is described by (14) and (18)-(19). First, the suboptimal estimator for S^2 will be described in detail; then the generalization to S^n will be discussed.

In our discussion of estimation on S^2, we will refer to a point on S^2 in terms of the Cartesian coordinates $x \triangleq (x_1, x_2, x_3)$ or the polar coordinates (θ, ϕ) (see the Appendix, in which harmonic analysis on S^n is summarized). According to the Appendix, the spherical harmonics $\{Y_{\ell m}\}$ (defined in (A.11)-(A.12)) are the eigenfunctions of the Laplacian Δ_{S^2} (defined in (A.10)), and all spherical harmonics of degree ℓ have the same eigenvalue $-\ell(\ell+1)$. Thus the assumed density approximation is a normal density on S^2 of the form (22), as discussed in the previous section:

$$p(\theta,\varphi,t) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\theta,\varphi)\, Y^*_{\ell m}(\eta(t),\lambda(t))e^{-\ell(\ell+1)\gamma(t)} . \tag{23}$$

where * denotes complex conjugate. In other words,

$$c_{\ell m}(t) \triangleq E^t[Y^*_{\ell m}(\theta(t),\phi(t))] = Y^*_{\ell m}(\eta(t),\lambda(t))e^{-\ell(\ell+1)\gamma(t)} . \tag{24}$$

In order to truncate the optimal estimator (18)-(19) after the $\hat{x}^{[N]}(t|t)$ equation using the assumed density (23), we must compute $E^t[x^{[N]}(t)x'(t)]$, or equivalently, $\hat{x}^{[N+1]}(t|t)$, in terms of $\hat{x}^{[p]}(t|t)$, $p = 1,2,\ldots,N$. However, if $\hat{x}(t|t)$ is known, so are $c_{10}(t)$ and $c_{11}(t)$. A simple computation [13], [22] then shows that $\{c_{N+1,m}, m = -(N+1),\ldots,N+1\}$ can be computed from

$$c_{N+1,m}(t) = Y^*_{N+1,m}(\eta(t),\lambda(t))e^{-(N+1)(N+2)\gamma(t)}$$

$$= (-1)^m \left[\frac{(N+1-m)!}{(N+1+m)!}\,\frac{2N+3}{4\pi}\right]^{\frac{1}{2}} P_{N+1,m}\left(\frac{c_{10}(t)}{(c_{10}^2(t)+2|c_{11}(t)|^2)^{\frac{1}{2}}}\right)$$

$$\cdot\left(\frac{c_{11}(t)}{c_{11}^*(t)}\right)^{m/2}\left[\frac{4\pi}{3}\left(c_{10}^2(t)+2|c_{11}(t)|^2\right)\right]^{\frac{1}{4}(N+1)(N+2)} \tag{25}$$

where $P_{\ell m}$ are the associated Legendre functions (see (A.11)-(A.12)). Finally, it is shown in [13] that there exists a nonsingular matrix T such that

$$Tx^{[N+1]} = \begin{bmatrix} Y_{N+1}(x) \\ x^{[N-1]} \end{bmatrix} \tag{26}$$

Thus $\hat{x}^{[N+1]}(t|t)$ can be computed from $\{c_{N+1,m}, -(N+1) \leq m \leq N+1\}$ and $\hat{x}^{[N-1]}(t|t)$. The optimal estimator (18) is truncated by substituting this approximation for $\hat{x}^{[N+1]}(t|t)$ into the equation for $\hat{x}^{[N]}(t|t)$. Notice that the entire procedure for truncating the optimal estimator can equivalently be performed on the infinite set of coupled equations for the generalized Fourier coefficients $c_{\ell m}(t)$, using the approximation (24).

We note that one can show that

$$\alpha(t) \triangleq \sqrt{\|\hat{x}(t|t)\|} \leq 1$$

and this quantity can be used as a measure of our confidence in our estimate. If $\hat{x}(t|t)$ satisfies the assumed density (23),

$$\alpha(t) = \|\hat{x}(t|t)\| = e^{-\gamma(t)} \tag{27}$$

so $\gamma = 0$ (zero "variance") implies $\alpha = 1$, and $\gamma = \infty$ (infinite "variance") implies $\alpha = 0$ (see [4] for the S^1 analog).

Example 1: Suppose that we truncate the optimal S^2 estimator (18) after N = 1--i.e., we approximate $\hat{x}^{[2]}(t|t)$ using the above approximation. Assume that Q(t) = I and $\{A_i, i = 1,2,3\}$ are given. Then the resulting suboptimal estimator is (for Q(t) = I)

$$d\hat{x}(t|t) = [A_0 + \tfrac{1}{2} \sum_{i=1}^{N} A_i^2]\, \hat{x}(t|t)dt$$

$$+ P(t)H'(t)R^{-1}(t)[dz_2(t) - H(t)\hat{x}(t|t)dt] \tag{28}$$

where the "covariance" matrix P(t) is given by

$$P_{ii}(t) = \hat{x}_i^2(t|t)(\tfrac{2}{3}\ \|\hat{x}(t|t)\| - 1) - \tfrac{1}{3}(\hat{x}_j^2(t|t) + \hat{x}_k^2(t|t))\|\hat{x}(t|t)\| + \tfrac{1}{3} \tag{29}$$

for $i \neq j$, $i \neq k$, $j \neq k$, and

$$P_{ij}(t) = \hat{x}_i(t|t)\hat{x}_j(t|t)(\|\hat{x}(t|t)\| - 1) \tag{30}$$

for $i \neq j$. It is shown in [37] that the matrix P(t) of (29)-(30) is positive semidefinite, and thus can be viewed as a covariance matrix. The results of Monte Carlo simulations to evaluate the performance of this estimator are presented in Section VII.

The extension to S^n of this technique for constructing suboptimal estimators is straightforward. The procedure uses the spherical harmonics on S^n. In polar coordinates, a point on S^n can be described by $(\theta_1, \theta_2, \ldots, \theta_{n-1}, \phi) \triangleq (\theta, \phi)$, where $0 \leq \theta_j \leq \pi$ and $0 \leq \phi \leq 2\pi$. Also, the spherical harmonics are denoted by

$$Y_{\ell,(m)}(\theta,\varphi) \triangleq Y_{\ell,m_1,\ldots,m_{n-1}}(\theta_1,\ldots,\theta_{n-1},\pm\varphi)$$

$$= e^{\pm i m_{n-1}\varphi} \prod_{k=0}^{n-2} (\sin\theta_{k+1})^{m_{k+1}}\, C_{m_k - m_{k+1}}^{m_{k+1}+\frac{1}{2}(n-k-1)}(\cos\theta_{k+1}) \tag{31}$$

where $\ell \geq m_1 \geq \ldots \geq m_{n-1} \geq 0$ and C_j^i are the Gegenbauer polynomials [33] (that is, the functions $Y_{\ell,(m)}$ are eigenfunctions of the Laplace-Beltrami operator with eigenvalue $-\ell(n+\ell-1)$). Hence the assumed density approximation on S^n is

$$p(\theta,\varphi,t) = \sum_{\ell,(m)} Y_{\ell,(m)}(\theta,\varphi)\, Y^*_{\ell,(m)}(\eta(t),\, \lambda(t))e^{-\ell(\ell+n-1)\gamma(t)} \tag{32}$$

That is, $c_{\ell,(m)}(t) \triangleq E^t[Y^*_{\ell,(m)}(\theta(t),\varphi(t))]$ is assumed to be

$$c_{\ell,(m)}(t) = Y^*_{\ell,(m)}(\eta(t),\lambda(t))e^{-\ell(\ell+n-1)\gamma(t)} . \tag{33}$$

The procedure for truncating the filter (18) is identical to the S^2 case. If $\hat{x}(t|t)$ is known, so are $c_{1,(m)}(t)$, and these can be used to compute $\gamma(t)$, $\eta(t)$, and $\lambda(t)$. Then $\{c_{N+1,(m)}(t)\}$ can be computed from (33), and $\hat{x}^{[N+1]}(t|t)$ can be computed from $\{c_{N+1,(m)}(t)\}$ and $\hat{x}^{[N-1]}(t|t)$. The estimator is truncated by substituting this approximate expression for $\hat{x}^{[N+1]}(t|t)$ into the equation (18) for $\hat{x}^{[N]}(t|t)$.

IV. ESTIMATION ON SO(n)

In this section we discuss the construction of suboptimal estimators for the SO(n) estimation problem (7)-(8). We will only consider the SO(3) problem; the results are extended to SO(n) in [13]. The concepts of harmonic analysis on SO(3) presented in the Appendix will be used extensively.

Consider the sequence $\{D^{\ell}, \ell = 0,1,\ldots\}$ of irreducible unitary representations of SO(3), as defined in (A.5)-(A.6). Theorem A.1 implies that, for fixed ℓ, the matrix elements $\{D^{\ell}_{mn}; -\ell \le m,n \le \ell\}$ are eigenfunctions of the Laplacian $\Delta_{SO(3)}$ (defined in (A.4)) with the same eigenvalue λ_{ℓ}; also, all eigenfunctions of the Laplacian can be written as linear combination of the $\{D^{\ell}_{mn}\}$. Hence, the assumed density which will be used to truncate the optimal estimator (16)-(17) is a normal density on SO(3) of the form (22):

$$p(R,t) = \sum_{\ell=0}^{\infty} \sum_{m,n=-\ell}^{\ell} D^{\ell}_{m,n}(R)\, D^{\ell}_{m,n}(\eta(t))^{*}\, e^{-\lambda_{\ell}\gamma(t)} \tag{34}$$

where R, $\eta(t) \in SO(3)$ and $\gamma(t)$ is a scalar. That is,

$$c^{\ell}_{mn}(t) \triangleq E^{t}[D^{\ell}_{mn}(\eta(t))^{*}] \tag{35}$$

is assumed to be

$$c^{\ell}_{mn}(t) = D^{\ell}_{mn}(\eta(t))^{*}\, e^{-\lambda_{\ell}\gamma(t)} . \tag{36}$$

The procedure for truncating the filter (16) is similar to the S^{n} case, although we make use of some additional concepts from representation theory. If $\hat{X}(t|t)$ is known, so are $\{c^{1}_{mn}(t); -1 \le m,n \le 1\}$, since D^{1} is equivalent to the self-representation of SO(3). Define the matrix $C^{\ell}(t)$ with elements $c^{\ell}_{mn}(t)$, $-\ell \le m,n \le \ell$; then

$$A(t) \triangleq \bar{C}^{1}(t)C^{1}(t) = [D^{1}(\eta(t))]'[D^{1}(\eta(t))]^{*}\, e^{-2\lambda_{1}\gamma(t)}$$

$$= I \cdot e^{-2\lambda_1 \gamma(t)} \tag{37}$$

since D^1 is unitary (here $\bar{C}$ is the hermitian transpose of C). Thus $\gamma(t)$ can be computed from

$$\gamma(t) = -\frac{1}{2\lambda_1} \log \left[\tfrac{1}{3} \operatorname{tr} A(t)\right] . \tag{38}$$

Then the elements of $\eta(t)$ can be computed from (36) and (38), since $D^1(\eta(t))$ is similar to $\eta(t)$. Once $\gamma(t)$ and $\eta(t)$ have been computed, $\{c^{N+1}_{mn}; -(N+1) \le m,n \le N+1\}$ are computed from the formula (36).

In order to truncate (16) after the N^{th} moment equation, we must approximate $E^t[X_v^{[N]}(t)h'(t)X(t)]$; however, this matrix consists of time-varying deterministic functions multiplying elements of $\hat{X}^{[N+1]}(t|t)$, so we will show how to approximate this matrix. The symmetrized Kronecker p^{th} power $X^{[p]}$ operating on the symmetric tensors $x^{[p]}$ such that $\|x^{[p]}\| = \|x\|^p = 1$ furnishes a representation of SO(3) which is reducible [26]. It is shown in [13] that there is a nonsingular matrix T such that

$$TX^{[p]}T^{-1} = \begin{bmatrix} D^p(X) & 0 \\ 0 & X^{[p-2]} \end{bmatrix} . \tag{39}$$

The matrix T is related to the Clebsch-Gordan coefficients [6], but T can also be computed by the method of Gantmacher [34, p. 160]. It is clear from the decomposition (39) that $\hat{X}^{[N+1]}(t|t)$ can be computed from $C^{N+1}(t)$ and $\hat{X}^{[N-1]}(t|t)$. The optimal estimator (16) is truncated by substituting this approximation into the equation for $\hat{X}^{[N]}(t|t)$.

We note here that, due to the decomposition (39), the estimation equations and the truncation procedure could have been expressed solely in terms of the irreducible representations $D^p(X(t))$. However, we have chosen to work with the $X^{[p]}$ equations primarily for ease of notation. For large N, the D^p equations would provide significant computational savings over the $X^{[p]}$ equations, as these are redundant; however, the practical implementation of this technique will probably be limited to small values of N.

V. SIMULATION RESULTS

As an illustration of the techniques presented in the previous sections, the first order filter (FOF) of Example 1(Section III) was evaluated by means of digital Monte Carlo simulations. It was compared to both the extended Kalman filter (EKF) [7] and the Gustafson-Speyer linear, minimum-variance quadrature filter (LQF) [24]. Identical noise sequences were used to allow direct comparisons.

The system considered was the S^2 system, i.e.,

$$dx(t) = F x(t)\, dt + \sum_{i=1}^{3} A_i x(t)\, dw_i(t) \tag{40}$$

$$dz(t) = x(t)\, dt + r^{1/2}\, dv(t) \tag{41}$$

where $F = \sum_{i=1}^{3} f_i A_i + \frac{1}{2} q \sum_{i=1}^{3} A_i^2$, and $\{A_i, i=1,2,3\}$ are the skew-symmetric matrices

$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{42}$$

Also, $w(t)$ has strength qI; and v has strength I. In this experiment, the nominal angular velocities $\{f_i, i=1,2,3\}$ were chosen to be 100.0, and q and r were varied.

For all three filters, the normalized estimate $\tilde{x}(t) = \hat{x}(t|t)/\|\hat{x}(t|t)\|$ was used. The filters have an identical structure for the approximate $\hat{x}$ equation:

$$d\hat{x}(t|t) = F\hat{x}(t|t) + \frac{1}{r} P(t)\, [dz(t) - \hat{x}(t|t)dt]\ . \tag{42}$$

However, for the FOF, $P(t)$ is given by the highly nonlinear memoryless equations (29)-(30). In the EKF, $P(t)$ satisfies the Riccati equation

$$\frac{d}{dt} P(t) = FP(t) + P(t)F' + qG(\hat{x}(t|t))G'(\hat{x}(t|t)) - \frac{1}{r} P(t)P'(t) \tag{43}$$

where

$$G(\hat{x}) = [A_1\hat{x},\ A_2\hat{x},\ A_3\hat{x}]\ . \tag{44}$$

Since the Riccati equation (43) is a function of $\hat{x}(t|t)$, the $P(t)$ calculation in the EKF requires <u>extensive</u> <u>on-line</u> <u>computation</u>, which represents a considerable burden.

In the LQF, P(t) is given as the solution of the coupled dynamic equations

$$\frac{d}{dt} P(t) = FP(t) + P(t)F' + \Delta(X(t),t) - \frac{1}{r} P(t)P'(t) \tag{45}$$

$$\frac{d}{dt} X(t) = FX(t) + X(t)F' + \Delta(X(t),t) \tag{46}$$

where $\Delta(X(t),t)$ is a diagonal matrix with ith component

$$\Delta(X(t),t)_i = q \sum_{k,\ell,m=1}^{3} (A_m)_{ik} (A_m)_{i\ell} X_{k\ell}(t) \tag{47}$$

Notice that these equations for P(t) and X(t) can be calculated off-line, but the LQF thus has a considerable storage requirement. Because P(t) in the FOF is only given by a memoryless nonlinearity, this filter requires considerably less storage than the LQF and less on-line computation than the EKF.

Our approach to the statistical analysis of the Monte Carlo simulations closely parallels that of Bucy and his associates [1], [3], [35]. The steady-state mean-squared error

$$\mu_2 = E[\|x(t) - \tilde{x}(t)\|^2] = \sum_{i=1}^{3} E[(x_i(t) - \tilde{x}_i(t))^2] \tag{48}$$

where $\tilde{x}_i(t)$ denotes the estimate of the ith component of the state $x_i(t)$, was used as the performance criterion. If $\{x^n\}$ and $\{\tilde{x}^n\}$, $i = 1,\ldots,N$, are sequences of independent realizations of x(t) and $\tilde{x}(t)$, respectively, then the statistic

$$\mu_2 = \frac{1}{N} \sum_{n=1}^{N} \|x^n - \tilde{x}^n\|^2 \tag{49}$$

is an approximation to μ_2 for sufficiently large N. In fact, by the Central Limit Theorem [36, p. 278], $\hat{\mu}_2$ is asymptotically normal with

$$E[\hat{\mu}_2] = \mu_2 \tag{50}$$

$$\text{var}\,[\hat{\mu}_2] = \frac{1}{N}\left\{\sum_{i=1}^{3} (\mu_4)_i + 2(\mu_4)_{12} + 2(\mu_4)_{13} + 2(\mu_4)_{23} - (\mu_2)^2\right\} \tag{51}$$

where $(\mu_4)_i \triangleq E[(x_i(t) - \tilde{x}_i(t))^4]$ and

$$(\mu_4)_{ij} = E[(x_i(t) - \tilde{x}_i(t))^2(x_j(t) - \tilde{x}_j(t))^2].$$

Thus, for large N, with probability 0.9974, the 3σ confidence interval is given by

$$|\mu_2 - \hat{\mu}_2| \leq 3\sqrt{\text{var}(\hat{\mu}_2)}\ , \tag{52}$$

or equivalently,

$$\Pr\left\{\frac{\hat{\mu}_2}{1 + 3\sqrt{\alpha}} \leq \mu_2 \leq \frac{\hat{\mu}_2}{1 - 3\sqrt{\alpha}}\right\} = 0.9974 \tag{53}$$

where

$$\alpha = \frac{\text{var}(\hat{\mu}_2)}{(\hat{\mu}_2)^2}\ . \tag{54}$$

In the Monte Carlo simulations, α was estimated from the samples (using sample means as in (49)), and approximate confidence intervals were thus computed.

In the experiment, 15 sample paths, each of which contained 1000 steps of length .001 seconds, were run in each simulation. The first 200 samples in each sample path were discarded to allow the transients to decay, so the remaining 800 samples represented steady-state. If all the steady-state errors were averaged as in (49), this would lead to 12000 samples of the steady-state error. However, as noted in [4], [24] and [35], adjacent errors in each path are correlated, so the effective Monte Carlo length is somewhere in the range between N=1200 and N=12000. The three standard deviation confidence intervals were calculated for both values of N.

The results of the simulations are presented in Table I. The 3σ confidence intervals I_1 (for N = 12000) and I_2 (for N = 1200) are shown. The results of this approximate statistical analysis of the Monte Carlo simulations indicate that, for this simple example, the FOF performs comparably to the LQF, and better than the EKF. The FOF seems to perform better in comparison to the other filters as q increases, due to the increasing dominance of the bilinear noise term in the system equation (40). These results are significant, due to the fact that the FOF designed here requires considerably less storage and computation than the other filters (no additional differential equations or storage for $P(t)$ are required).

		M.S. Error $\hat{u}_2$	I_1 (N=12000)	I_2 (N=1200)
$q = 0.01$	FOF	9.65	[9.41,9.91]	[8.92,10.52]
$r = 0.01$	EKF	10.21	[9.96,10.48]	[9.45,11.11]
	LQF	9.49	[9.26,9.75]	[8.78,10.34]
$q = 0.01$	FOF	11.69	[11.39,12.01]	[10.79,12.77]
$r = 1.00$	EKF	11.75	[11.44,12.07]	[10.84,12.82]
	LQF	11.69	[11.39,12.01]	[10.79,12.77]
$q = 1.00$	FOF	170.38	[165.26,175.82]	[155.19,188.87]
$r = 0.01$	EKF	193.75	[188.74,199.04]	[178.74,211.51]
	LQF	170.65	[165.58,176.03]	[155.60,188.92]

Table I. Monte Carlo M.S. Estimation Error(s)

($\times 10^{-3}$)

VIII. CONCLUSIONS

The state estimation problem for bilinear stochastic systems evolving on compact Lie groups and homogeneous space has been considered. The techniques of harmonic analysis on compact Lie groups have been applied to the design of suboptimal estimators for such systems. Monte Carlo simulations of a simple example indicate that a computationally simple filter designed by these methods performs favorably as compared to two other filters.

APPENDIX - HARMONIC ANALYSIS ON COMPACT LIE GROUPS

In this appendix we summarize some facts from the theory of harmonic analysis on Lie groups. For details of group representations see references [5,6,13,29,30].

Definition A.1: A finite-dimensional matrix representation of a compact Lie group G is a continuous homomorphism D which maps G into the group of nonsingular linear transformations on a finite dimensional vector space V. The representative ring of G is the ring generated over the field of complex numbers by the set of all continuous functions D_{ij} on G which are matrix elements of some irreducible representation D.

The Peter-Weyl theorem is the major result in harmonic analysis on compact Lie groups [5,6,29,30]. It gives the direct sum decomposition of $L_2(G)$, the space of square integrable functions with respect to the Haar measure dg:

$$L_2(G) = \oplus_{\alpha\in\Lambda} V_\alpha \tag{A.1}$$

where Λ is the set of equivalence classes of finite dimensional irreducible representations of G, and V_α denotes the vector space spanned by the $(n_\alpha)^2$ functions $\{D_{ij}^\alpha; i,j=1,\ldots,n_\alpha\}$.

The Laplace-Beltrami operator (Laplacian) on a Lie group or Riemannian homogeneous space is discussed in [5], [13] and [29]. Some examples will be presented here, but one very important result for our purposes is the following.

Theorem A.1 [5, p. 40], [29, p. 257]: Let G be a compact Lie group, and let V_α be defined as in equation (A.1). Then each function $\varphi \in V_\alpha$ is an eigenfunction of the Laplacian Δ, and all $\varphi \in V_\alpha$ have the same eigenvalue λ_α. Conversely, each eigenfunction φ of the Laplacian is an element of the representative ring.

Hence, harmonic analysis on a compact Lie group can be performed either in terms of the representative ring or the eigenfunctions of the Laplacian, since these two sets of functions are the same. In this paper we are primarily concerned with the application of these results to the special orthogonal group SO(n) and the n-sphere S^n.

The Lie group SO(n) is defined by

$$SO(n) = \{X \in \mathbb{R}^{n\times n} \mid X'X = I,\ \det X = +1\} \quad . \tag{A.2}$$

The theory of representations of SO(n) is discussed in [13], [26]; for this paper we need only consider SO(3). Any matrix R in SO(3) can be described in local coordinates in terms of the Euler angles ϕ, θ, ψ, domain $0 \leq \phi < 2\pi$, $0 \leq \theta \leq 2\pi$, $0 \leq \psi < 2\pi$ [6]. An element of SO(3) will thus be denoted by $R(\phi,\theta,\psi)$ or just (ϕ,θ,ψ).

In the Euler angle coordinates, the (unnormalized) Haar measure is

$$d\mu(\varphi,\theta,\Psi) = \sin\theta\, d\varphi\, d\theta\, d\Psi \tag{A.3}$$

and the corresponding Laplace-Beltrami operator is given by [31]

$$\Delta_{SO(3)} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2} - 2\cos\theta\frac{\partial^2}{\partial\varphi\partial\Psi} + \frac{\partial^2}{\partial\Psi^2}\right) \tag{A.4}$$

Talman [6] computes a sequence $D^\ell(\varphi,\theta,\Psi)$, $\ell = 0,1,\ldots,$ of unitary irreducible representations of SO(3); its matrix elements are given by

$$D^\ell_{mn}(\varphi,\theta,\Psi) = i^{m-n} e^{-im\varphi} d^\ell_{mn}(\theta) e^{-in\Psi} \tag{A.5}$$

where

$$d^{\ell}_{mn}(\theta) = \sum_t (-1)^t \frac{[(\ell+m)!(\ell-m)!(\ell+n)!(\ell-n)!]^{\frac{1}{2}}}{(\ell+m-t)!(t+n-m)!t!(\ell-n-t)!}$$

$$\cdot \cos^{2\ell+m-n-2t}\left(\frac{\theta}{2}\right)\sin^{2t+n-m}\left(\frac{\theta}{2}\right) \tag{A.6}$$

for $-\ell \leq m,n \leq \ell$. Here t is summed over all nonnegative integers such that the arguments of the factorial functions in (A.6) are nonnegative; i.e.,

$$m-n \leq t \leq \ell+m, \qquad 0 \leq t \leq \ell-n \ .$$

In fact, these are (up to equivalence) all of the irreducible representations of SO(3). The Peter-Weyl Theorem yields the decomposition

$$L_2(SO(3)) = \bigoplus_{\ell} V_{\ell} \tag{A.7}$$

where H_ℓ is the vector space spanned by the $(2\ell+1)^2$ functions $\{D^{\ell}_{mn};\ m,n = -\ell,\ldots,\ell\}$.

The n-sphere $S^n = \{x\varepsilon\mathbb{R}^{n+1} | x'x = 1\}$ is diffeomorphic to the homogeneous space SO(n)/SO(n-1). Harmonic analysis on S^n is studied in terms of the spherical harmonics [6,30,33]. The space H_ℓ of spherical harmonics of degree ℓ can be characterized as the eigenspace of the SO(n+1)-invariant Δ_{S^n} on S^n with eigenvalue $-\ell(n-1+\ell)$; other equivalent characterizations are given in [13].

In particular, we consider the 2-sphere S^2. Any point (x_1,x_2,x_3) on S^2 can be expressed in the polar coordinates (θ,ϕ), where $0 \leq \theta \leq \pi$, $0 \leq \phi < 2\pi$, by defining

$$x_1 = \cos\theta;\ x_2 = \sin\theta\cos\varphi;\ x_3 = \sin\theta\sin\varphi \ . \tag{A.8}$$

The Riemannian measure invariant under the action of SO(3) is

$$d\mu(\theta,\varphi) = \sin\theta\, d\theta d\varphi \ . \tag{A.9}$$

The corresponding invariant Laplace-Beltrami operator is

$$\Delta_{S^2} = \frac{1}{\sin\theta}\left[\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin\theta}\frac{\partial^2}{\partial\varphi^2}\right]. \tag{A.10}$$

The normalized spherical harmonics of degree ℓ on S^2 are defined by [6]

$$Y_{\ell m}(\theta,\varphi) = (-1)^m \left[\frac{(\ell-m)!}{(\ell+m)!}\,\frac{(2\ell+1)}{4\pi}\right]^{\frac{1}{2}} P_{\ell m}(\cos\theta)\, e^{im\varphi} \tag{A.11}$$

$$Y_{\ell,-m}(\theta,\varphi) = (-1)^m Y^*_{\ell m}(\theta,\varphi) \tag{A.12}$$

for $\ell=0,1,\ldots$ and $m = 0,1,\ldots,\ell$, where $P_{\ell m}(\cos\theta)$ are the associated Legendre functions and * denotes complex conjugate. Notice that $Y_{\ell m}$ is an eigenfunction of Δ_{S^2} with eigenvalue $-\ell(\ell+1)$.

REFERENCES

1. R.S. Bucy, C. Hecht, and K.D. Senne, "New Methods for Nonlinear Filtering," Revue Francaise d'Automatique, Informatique et de Recherche Operationnelle, February 1973, pp. 3-54.

2. R.S. Bucy and H. Youssef, "Fourier Realization of the Optimal Phase Demodulator," Fourth Symposium on Nonlinear Estimation and Its Applications, San Diego, CA., 1973.

3. J.J. Mallinckrodt, R.S. Bucy, and S.Y. Chang, Final Project Report for a Design Study for an Optimal Non-Linear Receiver/Demodulator, NASA Contract NAS 5-10789, Goddard Space Flight Center, Maryland, 1970.

4. A.S. Willsky, "Fourier Series and Estimation on the Circle with Applications to Synchronous Communication, Parts I, II," IEEE Trans. on Inf. Theory, Vol. IT-20, No. 5, September 1974, pp. 577-590.

5. E.M. Stein, Topics in Harmonic Analysis. Princeton: Princeton Univ. Press, 1970.

6. J.D. Talman, Special Functions--A Group Theoretic Approach. New York: Benjamin, 1968.

7. A.H. Jazwinski, Stochastic Processes and Filtering Theory. New York: Academic Press, 1970.

8. R.W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM J. Control, Vol. 10, No. 2, May 1972, pp. 265-284.

9. R.W. Brockett, "Lie Algebras and Lie Groups in Control Theory," in Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, eds., Reidel Publ. Co., The Netherlands, 1973.

10. R.W. Brockett, "Lie Theory and Control Systems on Spheres," SIAM J. Appl. Math., Vol. 25, No. 2, Sept., 1973, pp. 213-225.

11. A.S. Willsky, S.I. Marcus, and D.N. Martin, "On the Stochastic Stability of Linear Systems Containing Colored Multiplicative Noise," IEEE Trans. on Aut. Control, Vol. AC-20, Oct. 1975, pp. 711-713.

12. A.S. Willsky and S.I. Marcus, "Analysis of Bilinear Noise Models in Circuits and Devices," J. Franklin Inst., vol. 301, No. 1, January 1976, pp. 103-122.

13. S.I. Marcus, Estimation and Analysis of Nonlinear Stochastic Systems, Ph.D. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., June 1975; also M.I.T. Electronic Systems Laboratory Report No. ESL-R-601.

14. S.I. Marcus and A.S. Willsky, "Algebraic Structure and Finite Dimensional Nonlinear Estimation," SIAM Journal on Mathematical Analysis, to appear.

15. J.T. Lo, "Signal Detection on Lie Groups," in Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, eds., Reidel Pub. Co., The Netherlands, 1973.

16. J.M.C. Clark, "An Introduction to Stochastic Differential Equations on Manifolds," in Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, eds., Reidel Pub. Co., The Netherlands, 1973.

17. D.L. Elliott, "Diffusions on Manifolds Arising from Controllable Systems," in Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, eds., Reidel Pub. Co., The Netherlands, 1973.

18. J.T.H. Lo, "Exponential Fourier Densities and Optimal Estimation and Detection on the Circle," IEEE Trans. Inform. Theory, vol. IT-23, January 1977, pp. 110-116.

19. J.T.H. Lo and L.R. Eshleman, "Exponential Fourier Densities on S^2 and Optimal Estimation and Detection for Directional Processes," IEEE Trans. Inform. Theory, Vol. IT-23, May 1977, pp. 321-336.

20. J.T.H. Lo and L.R. Eshleman, "Exponential Fourier Densities on SO(3) and Optimal Estimation and Detection for Rotation Processes," in Proc. 1976 IEEE Conf. on Decision and Control, pp. 134-142.

21. A.S. Willsky and S.I. Marcus, "Estimation for Bilinear Stochastic Systems," in Variable Structure Systems with Applications to Economics and Biology, A. Ruberti and R.R. Mohler, eds., New York, Springer-Verlag, 1975, pp. 116-137.

22. S.I. Marcus and A.S. Willsky, "The Use of Spherical Harmonics in Suboptimal Estimator Design," in Proc. 1976 IEEE Conf. on Decision and Control, pp. 199-203.

23. I.Y. Bar-Itzhack, "Iterative Optimal Orthogonalization of the Strapdown Matrix," IEEE Trans. Aero. and Elec. Systems, Vol. AES-11, January 1975, pp. 30-37.

24. D.E. Gustafson and J.L. Speyer, "Linear Minimum Variance Filters Applied to Carrier Tracking," IEEE Trans. Aut. Control, Vol. AC-21, February 1976, pp. 65-73.

25. G.L. Blankenship, "Lie Theory and the Moment Stability Problem in Stochastic Differential Equations," Triennial World Congress of IFAC, Cambridge, MA., August 24-30, 1975.

26. F.D. Murnaghan, Theory of Group Representations. Baltimore: Johns Hopkins Press, 1938.

27. U. Grenander, Probabilities on Algebraic Structures. New York: John Wiley, 1963.

28. K. Yosida, Functional Analysis. New York: Springer-Verlag, 1974, pp. 398-403.

29. F. Warner, Foundations of Differential Manifolds and Lie Groups. Glenview, Ill.: Scott, Foresman and Co., 1971.

30. C.F. Dunkl and D.E. Ramirez, Topics in Harmonic Analysis. New York: Appleton-Century-Crofts, 1971.

31. R. Burridge, G.C. Papanicolaou, and D. McLaughlin, "The Solution by Descent of Certain Partial Differential Equations on the Sphere and Lobachevsky Plane," Comm. Pure Appl. Math., Vol. XXVI, 1973, pp. 105-129.

32. R.W. Brockett, Finite Dimensional Linear Systems. New York: John Wiley, 1970.

33. A. Erdelyi, W. Magnus, F. Oberhettinger, and F. Tricomi, Higher Transcendental Functions, Vol. II. New York: McGraw-Hill, 1953.

34. F.R. Gantmacher, The Theory of Matrices, Vol. I. New York: Chelsea, 1960.

35. R.S. Bucy, C. Hecht, and K.D. Senne, An Engineers' Guide to Building Nonlinear Filters, F.J. Seiler Research Lab. Report No. SRL-TR-72-0004, May 1972.

36. H. Bauer, Probability Theory and Elements of Measure Theory. New York: Holt, Rinehart, and Winston, 1972.

37. K. Hsu, The Evaluation of Several Suboptimal Filters for Bilinear Systems Evolving on Spheres, M.S. Thesis, Dept. of Electrical Engineering, University of Texas at Austin, Austin, Texas, May 1977.